HEAT AND GEOMETRY

JULIE ROWLETT

1. Lecture Notes 2018.07.13

Let $\Omega \subset \mathbb{R}^n$ be a bounded domain. By [vdbg 19, Theorem 1] and your exercise, there are constants $C, c > 0$ such that for all $x, y \in \Omega$ with $d(x, \partial\Omega) > \delta$ and $d(y, \partial\Omega) > \delta$ for any fixed $\delta > 0$ we have

$$|H_\Omega(t, x, y) - H_{\mathbb{R}^n}(t, x, y)| \leq Ce^{-c\delta/t}.$$

Above, H_Ω is the heat kernel on Ω, where we have assumed the Dirichlet boundary condition, and $H_{\mathbb{R}^n}$ is the heat kernel on $\mathbb{R}^n$,

$$H_{\mathbb{R}^n}(t, x, y) = (4\pi t)^{-n/2} e^{-|x-y|^2/(4t)}.$$

What exactly do we mean by "infinite speed of propagation?" First we shall prove the result for $\mathbb{R}^n$.

infrn **Theorem 1.** *Assume that $f \geq 0$ on $\mathbb{R}^n$, and there is some $\varepsilon > 0$ such that $f \geq \varepsilon$ on a set of positive measure. Moreover, assume that $f \in \mathcal{L}^2(\mathbb{R}^n)$. Let $U(t, x)$ be the solution to the heat equation on $\mathbb{R}^n$ with initial data given by f. Then for all $t > 0$ and for all $x \in \mathbb{R}^n$, $U(t, x) > 0$.*

Proof: We use the heat kernel to write out

$$U(t, x) = (4\pi t)^{-n/2} \int_{\mathbb{R}^n} e^{-|x-y|^2/(4t)} f(y) dy.$$

Since $f \geq 0$ on all of $\mathbb{R}^n$, we clearly have $U(t, x) \geq 0$ because everything else in the expression above is positive. Moreover, since $f \geq \varepsilon > 0$ on a set of positive measure, in fact

$$U(t, x) \geq (4\pi t)^{-n/2} \int_O \varepsilon e^{-|x-y|^2/4t} dy > 0,$$

where $O = \{y \in \mathbb{R}^n : f(y) \geq \varepsilon\}$.

Why is this called "infinite speed of propagation?" Well, the "heat" from the set on which the initial data f is positive travels *instantly* to everywhere else in $\mathbb{R}^n$. That is, as long as there is a little heat somewhere, and the initial data is non-negative, then that little bit of hit has reached every point in $\mathbb{R}^n$ after any length of time, no matter how short. It's like that South Park episode where Stan's money instantly disappears at the bank `https://www.youtube.com/watch?v=TGwZVGKG3Os` This works the other way, solve the heat equation with non-negative initial data which is positive somewhere, aaaand the heat's there. Instantly. Everywhere. I guess the physicists might be suspicious about this phenomenon?

We can prove the same fact for domains in $\mathbb{R}^n$ as well as for Riemannian manifolds. For the sake of simplicity we shall stick to the domains in $\mathbb{R}^n$ case. To prove the result we will use three ingredients:

(1) infinite speed of propagation for the heat equation on $\mathbb{R}^n$
(2) the maximum principle for solutions of the heat equation on bounded domains in $\mathbb{R}^n$
(3) the "principle of not feeling the boundary" (locality principle) made precise in [19, Theorem 1]

So at this point we need to prove the maximum principle.

Theorem 2 (Maximum principle (weak version)). *Let $\Omega \subset \mathbb{R}^n$ be a bounded (open) domain. Let $T > 0$, and define*

$$\Omega_T = \Omega \times [0, T), \quad \partial\Omega_T = \overline{\Omega} \times \{0\} \cup (\partial\Omega \times [0, T]).$$

If $u \in \mathcal{C}^2(\Omega_T) \cap \mathcal{C}^0(\overline{\Omega}_T)$ solves

$$\Xi u = 0$$

in Ω_T then

$$\max_{\Omega_T} u = \max_{\partial\Omega_T} u.$$

Proof: Define for $\varepsilon > 0$ the function $v(t, x) = u(t, x) - \varepsilon t$. Then

mp0
$$\Xi v = \Xi u - \varepsilon = -\varepsilon \implies \partial_t v - \sum \partial_i^2 v = -\varepsilon \implies \sum \partial_i^2 v = \varepsilon + \partial_t v. \tag{1.1}$$

Since $\overline{\Omega}_T$ is compact and v is continuous, it attains a maximum at some point (t_0, x_0). If this point is in the interior, then we must have $v_t = 0$ at this point and $\sum \partial_i^2 v \le 0$, which contradicts (1.1). Let us consider the case when v achieves its maximum on the interior of Ω at time T. Then, at this point we must have $v_t \ge 0$ and $\sum \partial_i^2 v \le 0$ which again contradicts (1.1). So, the maximum of v may occur in $\overline{\Omega}$ at $t = 0$ or on the boundary of Ω at $t = T$, or on the boundary of Ω for some other time. Well, that's precisely the set $\partial\Omega_T$. Since $v \to u$ uniformly as $\varepsilon \to 0$, we claim that we therefore obtain the result for u as well. To see this, assume for the sake of contradiction that u does not achieve its maximum on $\partial\Omega_T$. This means that u has its maximum at some point (t_0, x_0) which is not in $\partial\Omega_T$, and moreover, there is no point on $\partial\Omega_T$ at which u attains the value at (t_0, x_0). Thus $u(t, x) < u(t_0, x_0)$ for all points $(t, x) \in \partial\Omega_T$. Since $\partial\Omega_T$ is compact and u is continuous, u achieves its maximum on $\partial\Omega_T$. So, letting R be the maximum of u on $\partial\Omega_T$, we have

$$u(t_0, x_0) > R \implies \text{ letting } \delta = u(t_0, x_0) - R,$$

we have

$$u(t_0, x_0) = R + \delta > R + \frac{\delta}{2}.$$

So,

$$u(t_0, x_0) > u(t, x) + \frac{\delta}{2} \quad \forall (t, x) \in \partial\Omega_T.$$

So, we also have

$$v(t, x) = u(t, x) - \varepsilon t < u(t_0, x_0) - \frac{\delta}{2} - \varepsilon t \quad \forall (t, x) \in \partial\Omega_T.$$

Note that

$$v(t_0, x_0) = u(t_0, x_0) - \varepsilon t_0,$$

so when

$$\varepsilon = \frac{\delta}{2t_0} \implies v(t_0, x_0) = u(t_0, x_0) - \frac{\delta}{2} > v(t, x) \quad \forall (t, x) \in \partial\Omega_T.$$

Since (t_0, x_0) is not in $\partial\Omega_T$, we know that $t_0 \neq 0$, so we haven't done anything silly like dividing by zero. Since (t_0, x_0) is not in $\partial\Omega_T$, but v is supposed to achieve its maximum in $\partial\Omega_T$, we've hit a contradiction.

Corollary 3 (Minimum principle (weak version)). *Under the same hypotheses we also have*

$$\min_{\Omega_T} u = \min_{\partial\Omega_T} u.$$

Proof: Apply the maximum principle to $-u$ which satisfies the hypotheses of the theorem. This says that the maximum of $-u$ occurs on $\partial\Omega_T$. Since the maximum of $-u$ is equal to the minimum of u, we obtain the corollary.

We shall require a stronger assertion known as the strong maximum (minimum) principle. The proof relies on a certain mean value property satisfied by solutions of the heat equation. To state this mean value property, we define a **heat ball.**

Definition 4. A *heat ball* of radius r about the point (t, x) is the set

$$E(t, x; r) := \left\{ (y, s) \in \mathbb{R}^{n+1} : s \le t \text{ and } H_{\mathbb{R}^n}(t - s, x, y) \ge \frac{1}{r^n} \right\},$$

where of course

$$H_{\mathbb{R}^n}(t, x, y) = (4\pi t)^{-n/2} e^{-|x-y|^2/(4t)}.$$

Exercise 1. *Verify that*

$$E(t, x; r) = \left\{ (y, s) \in \mathbb{R}^{n+1} : t - \frac{r^2}{4\pi} \le s \le t, \quad |x - y|^2 \le 2n(t - s) \log\left(\frac{r^2}{4\pi(t - s)} \right) \right\},$$

and that

$$\partial E(t, x; r) = \left\{ (y, s) \in \mathbb{R}^{n+1} : t - \frac{r^2}{4\pi} \le s \le t, \quad |x - y|^2 = 2n(t - s) \log\left(\frac{r^2}{4\pi(t - s)} \right) \right\}.$$

This is why it's a bit weird to call it a heat ball because the point $(t, x) \in \partial E(t, x; r)$. On the other hand, please verify that $E(t, x; r) = (t, x) + E(0, 0; r)$. So, this may be one reason to call it a heat ball at the point (t, x). Got any ideas for a different name?

Solutions of the heat equation satisfy a certain mean value property given in terms of the integral over a heat ball. Note that it doesn't actually matter here what boundary condition we take.

Theorem 5 (Mean value property). *Let $\Omega \subset \mathbb{R}^n$ be a bounded (open) domain. Let $T > 0$, and define*

$$\Omega_T = \Omega \times [0, T), \quad \partial\Omega_T = \overline{\Omega} \times \{0\} \cup (\partial\Omega \times [0, T]).$$

If $u \in \mathcal{C}^2(\Omega_T) \cap \mathcal{C}^0(\overline{\Omega}_T)$ solves

$$\Xi u = 0$$

in Ω_T *then for each* $E(x,t;r) \subset \Omega_T$,

$$u(t,x) = \frac{1}{4r^n} \int_{E(x,t;r)} u(y,s) \frac{|x-y|^2}{(t-s)^2} dyds.$$

Proof: As you've shown in the exercise, changing the center of the heat ball just translates it. So, we may as well slide Ω_T around so that the point $(x,t) = (0,0)$. Observe that for the purpose of being a solution to the heat equation as well as in the definition of the heat ball, time can be positive of negative, no problem.

So, we consider

$$\phi(r) := \frac{1}{r^n} \int_{E(r)} u(y,s) \frac{|y|^2}{s^2} dyds = \int_{E(1)} u(ry, r^2 s) \frac{|y|^2}{s^2} dyds,$$

where $E(r) = E(0,0;r)$. We wish to show that this is a constant function of r. It's going to be convenient to switch coordinates to make r into 1 (and quite possibly switch back. So, let us define z and τ via

$$y = rz, \quad s = r^2 \tau.$$

Then, in the definition of the heat ball,

$$-\frac{r^2}{4\pi} \le s \le 0 \implies -\frac{1}{4\pi} \le \tau \le 0,$$

and

$$|rz|^2 = |y|^2 \le 2n(-s) \log\left(\frac{r^2}{4\pi(-s)}\right) \implies |z|^2 \le 2n\left(-\frac{s}{r^2}\right) \log\left(\frac{r^2}{4\pi(-r^2\tau)}\right) = 2n(-\tau)\log\left(\frac{1}{4\pi(-\tau)}\right).$$

So, in the coordinates (z,τ) the heat ball $E(r)$ because $E(1)$. Since $y = rz \implies dy = r^n dz$, and similarly $ds = r^2 d\tau$, we have

$$\phi(r) = r^{-n} \int_{E(1)} u(rz, r^2\tau) \frac{|rz|^2}{(r^2\tau)^2} r^n dz r^2 d\tau = \int_{E(1)} u(rz, r^2\tau) \frac{|z|^2}{\tau^2} dz d\tau.$$

By the regularity of u on a compact set, we can exchange integration and differentiation to compute

$$\phi'(r) = \int_{E(1)} \left(\nabla_z u \cdot z \frac{|z|^2}{\tau^2} + 2r u_\tau \frac{|z|^2}{\tau} \right) dz d\tau.$$

Doing a similar change of variables this is

$$\frac{1}{r^{n+1}} \int_{E(r)} \nabla_y u(y,s) \cdot y \frac{|y|^2}{s^2} dyds + \frac{2}{r^{n+1}} \int_{E(r)} u_s(y,s) \frac{|y|^2}{s} dyds$$
$$= A + B.$$

We're going to want to do some integration by parts tricking around, so we introduce

$$\psi(y,s) = \log(r^n H_{\mathbb{R}^n}(-y,-s)) = \log\left(\left(\frac{r^2}{-4\pi s}\right)^{n/2} e^{|y|^2/(4s)} \right) = n \log r - \frac{n}{2} \log(-4\pi s) + \frac{|y|^2}{4s}.$$

Then note that

$$\nabla_y \psi(y,s) = \frac{y}{2s}, \quad \psi = 0 \text{ on } \partial E(r).$$

So, the part

$$\frac{|y|^2}{s} = 2\nabla_y \psi(y,s) \cdot y.$$

Hence
$$B = \frac{4}{r^{n+1}} \int_{E(r)} u_s(y,s) \nabla_y \psi(y,s) \cdot y dyds.$$
To use integration by parts we take the divergence of $u_s \psi y$, so that
$$B = \frac{4}{r^{n+1}} \int_{E(r)} \left(\operatorname{div}\left(u_s(y,s)\psi(y,s)y\right) - \psi(y,s)\nabla_y u_s(y,s)\cdot y - n\psi(y,s)u_s(y,s)\right) dyds.$$
By the divergence theorem and the vanishing of ψ on the boundary of $E(r)$, the div term dies, so we have
$$B = -\frac{4n}{r^{n+1}} \int_{E(r)} \psi(y,s)u_s(y,s)dyds - \frac{4}{r^{n+1}} \int_{E(r)} \psi(y,s)\partial_s\left(\nabla_y u(y,s)\cdot y\right) dyds.$$
To handle the second term, we do integration by parts in the s variable to obtain
$$\begin{aligned} -\int_{E(r)} \psi(y,s)\partial_s\left(\nabla_y u(y,s)\cdot y\right) dyds &= \int_{E(r)} \partial_s \psi(y,s) \nabla_y u(y,s)\cdot y dyds \\ &= \int_{E(r)} \left(-\frac{n}{2s} - \frac{|y|^2}{4s^2}\right) \nabla u(y,s)\cdot y dyds. \end{aligned}$$
Now we put the A and B terms together...
$$\begin{aligned} A + B = \frac{1}{r^{n+1}} \int_{E(r)} \nabla_y u(y,s)\cdot y \frac{|y|^2}{s^2} dyds - \frac{4n}{r^{n+1}} \int_{E(r)} \psi(y,s)u_s(y,s)dyds \\ -\frac{4}{r^{n+1}} \int_{E(r)} \left(\frac{n}{2s} + \frac{|y|^2}{4s^2}\right) \nabla u(y,s)\cdot y dyds. \end{aligned}$$
So we get some nice cancellation in the first and last terms,
$$A + B = -\frac{4n}{r^{n+1}} \int_{E(r)} \psi(y,s)u_s(y,s)dyds - \frac{2n}{r^{n+1}} \int_{E(r)} \frac{\nabla u(y,s)\cdot y}{s} dyds.$$
Here's where we finally use that u solves the heat equation! Our heat equation is $\partial_t u - \operatorname{div} \nabla u = 0$. So, we can write
$$-\frac{4n}{r^{n+1}} \int_{E(r)} \psi(y,s)u_s(y,s)dyds = -\frac{4n}{r^{n+1}} \int_{E(r)} \psi(y,s) \operatorname{div} \nabla u = \frac{4n}{r^{n+1}} \int_{E(r)} \nabla\psi(y,s)\cdot\nabla u(y,s)dyds,$$
where we have used integration by parts and killed the boundary term using ψ. Hence,
$$A + B = \frac{4n}{r^{n+1}} \int_{E(r)} \nabla\psi(y,s)\cdot\nabla u(y,s)dyds - \frac{2n}{r^{n+1}} \int_{E(r)} \frac{\nabla u(y,s)\cdot y}{s} dyds.$$
Recalling that $\nabla\psi(y,s) = \frac{y}{2s}$,
$$A + B = \frac{4n}{r^{n+1}} \int_{E(r)} \frac{\nabla u(y,s)\cdot y}{2s} dyds - \frac{2n}{r^{n+1}} \int_{E(r)} \frac{\nabla u(y,s)\cdot y}{s} dyds = 0.$$
Super! So we have proven that $\phi(r)$ is a constant function of r. To determine its value, we let $r \downarrow 0$ and use the regularity of u to conclude
$$\lim_{r\to 0} \phi(r) = \lim_{r\to 0} \int_{E(1)} u(ry, r^2 s) \frac{|y|^2}{s^2} dyds = u(0,0) \int_{E(1)} \frac{|y|^2}{s^2} dyds.$$

Exercise 2. *Read and verify the steps in the calculation of the heat ball integral contained in the supplementary notes for this lecture.*

With the exercise, it turns out that the integral in question evaluates to 4. It's rather amazing. So, we get

$$\phi(r) \equiv 4u(0,0) \quad \forall r,$$

which completes the proof.

Theorem 6 (Maximum principle (strong version))**.** *Under the same hypotheses, if u achieves its maximum or minimum at a point which is not in $\partial\Omega_T$, then u is constant.*

Proof: Assume that there is an interior point at which u achieves its maximum. By the mean value property, for sufficiently small r such that the heat ball sits inside Ω_T, then u must be constant on that ball. Since Ω is connected and compact, we obtain that for any other interior point we can start at the maximum point (x_0, t_0) and make little heat balls starting there and connecting up to any other point (y_0, s_0). The function is constant in each of the balls, so finally we get the same value at all points in Ω_T. The same argument works equally well if we assumed that u achieves its minimum at an interior point of Ω_T.

Theorem 7. *Assume that $f \geq 0$ on all of Ω, and that $f \in \mathcal{L}^2_0(\Omega) \cap \mathcal{C}^0$. Here, $\mathcal{L}^2_0$ is the $\mathcal{L}^2$ closure of smooth, compactly supported functions in Ω. Assume that for some $\varepsilon > 0$ there is a set of positive measure on which $f \geq \varepsilon$. Let $u(t,x)$ be the solution to the initial value problem for the heat equation on Ω with initial data given by f. Then for any $x \in \Omega$ (that is x in the interior of Ω not on $\partial\Omega$), and for any $t > 0$,*

$$u(t,x) > 0.$$

Proof: First let us fix the point $x \in \Omega$. Let $\delta = d(x, \partial\Omega) > 0$. Extend the initial data, f to be identically zero on $\mathbb{R}^n \setminus \Omega$. Let

$$u(t,x) = \int_\Omega H_\Omega(t,x,y) f(y) dy,$$

and

$$U(t,x) = \int_{\mathbb{R}^n} H_{\mathbb{R}^n}(t,x,y) f(y) dy = \int_\Omega H_{\mathbb{R}^n}(t,x,y) f(y) dy.$$

Then we simply estimate

$$|u(t,x) - U(t,x)| = \left| \int_\Omega f(y) \left(H_\Omega(t,x,y) - H_{\mathbb{R}^n}(t,x,y) \right) dy \right|.$$

We split the integral into the integral over $\Omega_0 = \{y \in \Omega : d(y, \partial\Omega) \geq \frac{\delta}{2}\}$ and $\Omega_u = \{y \in \Omega : d(y, \partial\Omega) < \frac{\delta}{2}\}$. Then, we note that since

$$d(x, \partial\Omega) = \delta \implies d(x,y) \geq \frac{\delta}{2} \forall y \in \Omega_u.$$

Hence we have for both heat kernels, by the Cauchy-Schwarz inequalit

$$\int_{\Omega_u} |f(y)||H_*(t,x,y)| dy \leq ||f||_{\mathcal{L}^2(\Omega)} \sqrt{\int_{\Omega_u} |H_*(t,x,y)|^2 dy} \lesssim e^{-c\delta/t}, \quad * = \Omega, \mathbb{R}^n.$$

This is due to the off-diagonal decay of the heat kernels. Inside Ω_0, the points (x,y) are at least a distance $\frac{\delta}{2}$ from the boundary, so we can apply [19, Theorem 1] and the Cauchy-Schwarz inequality to estimate

$$\left| \int_{\Omega_0} f(y) \left(H_\Omega(t,x,y) - H_{\mathbb{R}^n}(t,x,y) \right) dy \right| \le ||f||_{\mathcal{L}^2(\Omega)} \sqrt{\int_{\Omega_0} C e^{-c\delta/t} dy}$$
$$\le ||f||_{\mathcal{L}^2(\Omega)} \sqrt{\mathrm{vol}(\Omega)} e^{-c\delta/t}.$$

We therefore obtain the estimate for some other constants $c, C > 0$

$$|u(t,x) - U(t,x)| \le C e^{-c/t}.$$

Taking $t > 0$ sufficiently small, we obtain

$$u(t,x) > \frac{U(t,x)}{2}.$$

By Theorem 1,

$$U(t,x) > 0 \quad \forall t > 0, \quad \forall x \in \mathbb{R}^n.$$

Hence, we have obtained that $u(t,x) > 0$ for sufficiently small t. By the maximum and minimum principles, $u(t,x) \ge 0$ for all x and t. If, for the sake of contradiction we were to have $u(T,x) = 0$ for some x inside Ω and some $T > 0$, then this would be a minimum point of u since $u \ge 0$. By the strong maximum principle, this would imply that u is in fact constant, so we would need to have $u \equiv 0$. This contradicts the fact that at $t = 0$ u is positive on a set of positive measure.

So, we see that in this case as well the "heat" from the set on which the initial data f is positive travels *instantly* to everywhere else in Ω. Hence the first of the "dangers" of heat: infinite speed of propagation.

1.1. **Microscopic perspective.** Let's now shift focus to the microscopic perspective. Since $\mathbb{R}^n$ is the n-fold product of $\mathbb{R}$, it suffices to work with $\mathbb{R}$ because we can just take the product of several copies to obtain $\mathbb{R}^n$. Let's consider a particle moving randomly along one dimension. In many theories of physics, one makes the assumption that at the smallest scales, space and time are discrete. So, let's do this and assume that the particle moves either one unit to the left or one unit to the right, with equal probabilities. Let x be the starting point in $\mathbb{R}$ of our particle. Then, after n steps,

$$S_n = x + \sum_{i=1}^n X_i, \quad X_i \in \{\pm 1\} \text{ are independent random variables.}$$

For the sake of simplicity, let us assume $x = 0$. The probabilities

$$P[X_j = 1] = P[X_j = -1] = \frac{1}{2}.$$

The expectation (or expected value) is by definition the sum over all possible values, z, the random variable, X can assume, of the value, z, times the probability that $X = z$,

$$E[X] = \sum_z z P[X = z] \implies E[S_n] = \sum_{j=1}^n P[X_j = 1] - P[X_j = -1] = 0.$$

Since the variables X_j are all independent, we can write $E[S_n]$ as above. Next, we compute $E[S_n^2]$,

$$E[S_n^2] = E[\left(\sum_{j=1}^n X_j\right)^2] = E[\sum X_j \sum X_k] = \sum\sum E[X_j X_k],$$

because $E[X+Y] = E[X] + E[Y]$ whenever X and Y are independent. So, since $X_j^2 = 1$ for all j,

$$E[S_n^2] = n + \sum_{j \neq k} E[X_j X_k].$$

The second part

$$\sum_{j \neq k} E[X_j X_k] = \sum_{X_j X_k = 1} P[X_j X_k = 1] - \sum_{X_j X_k = -1} P[X_j X_k = -1] = 0.$$

So,

$$E[S_n^2] = n.$$

The variance of a random variable,

$$\text{Var}[X] = E[(X - E[X])^2] = E[X^2] - E[X]^2.$$

So, in our case

$$\text{Var}[S_n] = E[S_n^2] - E[S_n]^2 = n.$$

We note that whenever the variables are independent, the variance is additive. That is

$$\text{Var}[X+Y] = \text{Var}[X] + \text{Var}[Y].$$

It doesn't obviously look like this is true, but if we write it out,

$$\text{Var}[X+Y] = E[(X+Y)^2] - E[X+Y]^2 = E[X^2 + 2XY + Y^2] - (E[X] + E[Y])^2$$
$$= E[X^2] + 2E[XY] + E[Y^2] - E[X]^2 - 2E[X]E[Y] - E[Y]^2.$$

By definition of expected value, $E[XY] = E[X]E[Y]$. So, the nice cancellation occurs, leaving

$$E[X^2] + E[Y^2] - E[X]^2 - E[Y]^2 = \text{Var}[X] + \text{Var}[Y].$$

What is the probability that our random walk returns to the starting point, 0. If n is odd, then S_n is odd, because $X_j + X_k$ for any j, k is always even. So, if the total number n of steps is odd, then the sum $S_n = X_1 + \ldots + X_{n-1} + X_n$ is equal to some even number plus X_n. Of course, X_n is odd. So the sum S_n is odd. However 0 is even. So we never reach the starting point with an odd number of steps. We also see by this logic that X_n is even whenever n is even. So, we can only return to zero after an even number of steps. Moreover, need to take n steps to the right (+1) and n steps to the left. How many ways are there to do this?

$$\binom{2n}{n}.$$

The probability if each choice is 2^{-2n}. Hence, the probability

$$P[S_{2n} = 0] = \binom{2n}{n} 2^{-2n} = \frac{(2n)!}{n!n!} 2^{-2n}.$$

More generally,

$$P[S_{2n} = 2j] = \binom{2n}{n+j} 2^{-2n},$$

because we need $n+j$ steps to the right and $n-j$ steps to the left to end up at $2j$. At this point it is perhaps curious why we are just looking at $2n$ and $2j$, that is even numbers of steps and even positions. The reason is that it will simplify calculations which are already unwieldy. In the end, no generality is lost, because we could just (1) cut the step-size (discretization) in half. Then every position in our full-step size universe is an even number of half steps. So, we aren't losing any generality by just looking at positions $2j$.

Since n will be getting large and we have binomial coefficients and factorials running around, it is useful to recall

Theorem 8 (Stirling's Formula with remainder estimate)**.** *For large n,*

$$n! = n^n \sqrt{2\pi n} e^{-n} \left(1 + \mathcal{O}\left(\frac{1}{n}\right)\right).$$

Let us study the asymptotic behavior of $P[S_{2n} = 2j]$ as $n \to \infty$. We have computed this to be

$$\binom{2n}{n+j} 2^{-2n} = \frac{(2n)!}{(n+j)!(n-j)!} 2^{-2n}.$$

By Stirling's formula, this is

$$\frac{2^{-2n}(2n)^{2n}\sqrt{2\pi(2n)}e^{-2n}(1+\mathcal{O}(1/n))}{((n-j)^{n-j+\frac{1}{2}}\sqrt{2\pi}e^{-(n-j)})(1+\mathcal{O}(1/(n-j)))(n+j)^{n+j+\frac{1}{2}}\sqrt{2\pi}e^{-(n+j)}(1+\mathcal{O}(1/(n+j))}.$$

To handle these remainders, we need to make some assumption about the relationship between j and n. The position is $2j$, achieved after $2n$ steps. So, we identify $2j$ with position and $2n$ with time. Inspired by the heat kernel in $\mathbb{R}$, let us assume that $j \sim \sqrt{n}$. Simplifying our expression and considering the remainder term all together, we have

$$\begin{aligned}
&\frac{n^{2n+\frac{1}{2}}}{\sqrt{\pi}(n+j)^{n+j+\frac{1}{2}}(n-j)^{n-j+\frac{1}{2}}} (1+\mathcal{O}(1/n)) \\
&= \frac{n^{-1/2}}{\sqrt{\pi}(1-\frac{j}{n})^{n-j+\frac{1}{2}}(1+\frac{j}{n})^{n+j+\frac{1}{2}}} (1+\mathcal{O}(1/n)) \\
&= \frac{1}{\sqrt{\pi}} n^{-\frac{1}{2}} \left(1-\frac{j}{n}\right)^{j-\frac{1}{2}} \left(1+\frac{j}{n}\right)^{-j-\frac{1}{2}} \left(1-\left(\frac{j}{n}\right)^2\right)^{-n} (1+\mathcal{O}(1/n)).
\end{aligned}$$

Now since we are taking j on the scale of $\sqrt{n}$, let us write $j = c\sqrt{n}$ for some bounded (possibly changing/variable) positive number c which is bounded above and below by fixed positive constants. Then our expression becomes

$$\frac{1}{\sqrt{\pi n}} \left(\left(1+\frac{-c}{\sqrt{n}}\right)^{\sqrt{n}}\right)^{c} \left(\left(1+\frac{c}{\sqrt{n}}\right)^{\sqrt{n}}\right)^{-c} \left(1-\frac{c^2}{n}\right)^{-n} \left(1-\frac{c^2}{n}\right)^{-\frac{1}{2}} (1+\mathcal{O}(1/n)).$$

We need a few more remainder estimates.

Proposition 9. *For $x \in \mathbb{R}$ fixed, we have for large n,*

$$\left(1+\frac{x}{n}\right)^n = e^x\left(1+\mathcal{O}(1/n)\right), \quad n \to \infty,$$

and

$$\left(\frac{1-\frac{x}{\sqrt{n}}}{1+\frac{x}{\sqrt{n}}}\right)^{\sqrt{n}} = e^{-2x}\left(1+\mathcal{O}(1/n)\right).$$

Proof: The proof is left as an exercise, but with a few hints.

(1) To prove the first statement, start by changing variables to $n = 1/s$, and consider
$$(1+sx)^{1/s}.$$
(2) Do a Taylor series expansion in terms of the variable s about some real number $s = \varepsilon > 0$.
(3) Show that the first term will converge to e^x.
(4) Estimate the remainder.
(5) For the second statement, proceed similarly. First replace $\sqrt{n}$ with n and show that the remainder is like n^{-2}.
(6) Do the same kind of s and $1/n$ swap as before, so you're considering
$$\left(\frac{1-sx}{1+sx}\right)^{\frac{1}{s}}.$$
(7) Do the same Taylor expansion procedure.
(8) Show that the leading term is e^{2x} and the remainder is $\mathcal{O}(s^2)$.

We proceed using our remainder estimates. Our expression is now

$$\frac{1}{\sqrt{\pi n}}e^{-2c^2}e^{c^2} + \mathcal{O}(1/n) = \frac{1}{\sqrt{\pi n}}e^{-k^2/n} + \mathcal{O}(1/n).$$

Now we see that for $a < b$,

$$P[a\sqrt{2n} \le S_{2n} \le b\sqrt{2n}] = \sum_{a\sqrt{2n}\le 2j \le b\sqrt{2n}} \left(\frac{e^{-j^2/n}}{\sqrt{n\pi}} + \mathcal{O}(1/n)\right).$$

Note that the inequality is equivalent to

$$a \le j\frac{\sqrt{2}}{\sqrt{n}} \le b.$$

So, we recognize a Riemann sum with step size $\frac{\sqrt{2}}{\sqrt{n}}$ for the function $f(x) = e^{-x^2/2}$,

$$\sum_{a\sqrt{2n}\le 2j \le b\sqrt{2n}} \left(\frac{e^{-j^2/n}}{\sqrt{n\pi}} + \mathcal{O}(1/n)\right) = \sum_{a\sqrt{2n}\le 2j \le b\sqrt{2n}} f\left(j\sqrt{\frac{2}{n}}\right)\sqrt{\frac{2}{n}}\frac{1}{\sqrt{2\pi}} + \sum_{a\sqrt{2n}\le 2j \le b\sqrt{2n}} \mathcal{O}(1/n).$$

The terms in the sum is on the order of $\sqrt{n}$, so the remainder term becomes

$$\mathcal{O}(1/\sqrt{n}).$$

Letting $n \to \infty$, the Riemann sum converges to the integral

$$\int_a^b \frac{e^{-x^2/2}}{\sqrt{2\pi}}dx.$$

Now if we started at a different point, say y, we would just get a shift in the integral. I leave this as an exercise. So, if the random walk starts at y, then the probability of being between a and b is tending to

$$\int_a^b \frac{e^{-(x-y)^2/2}}{\sqrt{2\pi}} dx.$$

Our discrete random walk tends towards a continuous random walk $\{W_t\}_{t\geq 0}$ where are random variables indexed by t. Let $d_t = \frac{1}{n}$ be time increments. Then, $W_{kd_t} \approx d_x S_k$, where d_x is the length of a step, and S_k is a random walk with step length 1. In other words, $d_x S_k$ is a discrete random walk with step-length d_x. To figure out what d_x should be, we normalize so that $E[W_1^2] = 1$. Then

$$E[(d_x S_n)^2] = d_x^2 E[S_n^2] = d_x^2 n \implies d_x = \frac{1}{\sqrt{n}}, \quad d_x = \sqrt{d_t}.$$

Thus the step-size $d_x = \sqrt{d_t}$. That is consistent with our calculations where $2j$ was the position and $2n$ was the number of steps, and $j \sim \sqrt{n}$. So, if $t = jd_t = \frac{j}{n}$ then

$$W_t = W_{jd_t} = W_{j/n} \approx d_x S_j = \frac{S_j}{\sqrt{n}} = \frac{S_j}{\sqrt{j}}\sqrt{\frac{j}{n}} = \sqrt{t}\frac{S_j}{\sqrt{j}}, \quad t = \frac{j}{n}.$$

What's going on here? The idea is that the step sizes are getting smaller and smaller, and the time increments are getting smaller and smaller, to make the discrete random walk tend towards the continuous random walk, W_t. So, the term $\frac{S_j}{\sqrt{j}}$ just corresponds to our discrete random walk where the position is like j, and the number of steps is like $n = j^2$. In other words, we are specifying the rate at which the time steps and the position steps are shrinking, that is the rate at which both j and n are tending to infinity.

We proved that when $j \sim \sqrt{n}$, then

$$P[S_j/\sqrt{j} \in [a,b]] = \int_a^b \frac{e^{-(x-y)^2/2}}{\sqrt{2\pi}} dx, \quad y \text{ is the starting point of the random walk.}$$

The integrand is known as the density function for the normal distribution with mean y and variance 1. The normal distribution with mean y and variance σ^2 has the density function

$$e^{-(x-y)^2/(2\sigma^2)}(2\pi\sigma^2)^{-1/2}.$$

Since we have $S_j/\sqrt{j}$ scaled by $\sqrt{t}$ in front, and the variance scales by the square, $\sqrt{t}S_j/\sqrt{j}$ will tend to have density function

$$e^{-(x-y)^2/2t}(2\pi t)^{-1/2}.$$

Thus for the continuous random walk starting at the point y, the probability of being in the interval $[a,b]$ at time t is

$$\int_a^b e^{-(x-y)^2/2t}(2\pi t)^{-1/2} dx.$$

Exercise 3. *Show that if we change our Laplacian by a factor of $\frac{1}{2}$, the heat kernel for the operator $\partial_t + \frac{1}{2}\Delta$ on $\mathbb{R}$ is $e^{-(x-y)^2/2t}(2\pi t)^{-1/2}$.*

So, the macroscopic flow of heat is driven by the microscopic random motion of particles.

2. Concluding remarks: heat and geometry

Let us return now to the motivation for this lecture series: the connection between geometry and heat.

2.1. Short time asymptotic expansion of the heat trace. The short time asymptotic expansion of the heat trace on a smooth, closed manifold (M, g) of dimension n with no boundary is of the form

$$t^{-n/2} \sum_{j \geq 0} a_j t^j,$$

where each

$$a_j = \int_M u_j(z, z) dz,$$

are integrals of universal polynomials in the curvature tensor and its covariant derivatives. This means that it is the same polynomial for any manifold. If we pass to the case where (M, g) has *smooth* boundary, which includes the case that M is a domain in $\mathbb{R}^n$ and g is the Euclidean metric, then there are additional terms in the expansion coming from the curvature of the boundary. The expression is

$$\sum_{j \geq 0} a_j t^{(j-n)/2}.$$

In certain singular settings, there is also such an expansion. When such an expansion exists, it can be used to define the zeta-regularized determinant of the Laplacian. This is defined through

$$\zeta(s) = \sum_{\lambda_k > 0} \lambda_k^{-s},$$

where λ_k are the eigenvalues of the Laplacian on M.

Exercise 4. *Prove that*

$$\zeta(s)\Gamma(s) = \int_0^\infty t^{s-1} \sum_{\lambda_k > 0} e^{-\lambda_k t}.$$

By Weyl's law, $\zeta(s)$ is well-defined and holomorphic for $\Re(s) > \frac{n}{2}$. Using the short time asymptotic expansion of the heat trace, we can prove that in fact the zeta function extends and is holomorphic around $s = 0$. In this way one can define

$$\det(\Delta) := e^{-\zeta'(0)}.$$

Exercise 5. *Show that were things to converge, we would have*

$$\det(\Delta) = \prod_{\lambda_k > 0} \lambda_k = e^{-\zeta'(0)}.$$

The determinant of the Laplacian is quite interesting to physicists [8] and mathematicians [16, 17]. In the physics literature [8], the determinant is somehow connected to Feynmann path integrals; don't ask me how this works! In the mathematics literature, it has been used to prove that the set of isospectral Riemannian metrics on a smooth surface is compact in the smooth topology [16, 17]. Interestingly, in that same work, they showed that the determinant is monotone along the Ricci flow. It's pretty rad.

Other applications of this expansion show that "one can hear" certain geometric features. For example, Watanabe [21, 22] proved that "one can hear" certain ellipse-like domains, in the sense that if any other domain is isospectral, then it is in fact the same. There are other results like "one can hear the corners of a drum" [11, 15], and "one can hear the regular n-gon" [10].

The short time asymptotic expansion of the heat trace is "local" in the sense that we can understand it by understanding it in each neighborhood without considering other neighborhoods which are far away. On the other hand, if we look at the heat kernel at very large times, everything mixes together, and we can no longer work locally. It becomes a "global" object. The long-time behavior of the heat kernel is connected to the topology [14].

In conclusion, the heat kernel is intimately linked to the curvature and geometry of the manifold or domain on which it lives. Many connections between the heat kernel and geometry have been discovered, but there is still much we do not know, and perhaps you shall contribute!

References

[1] A. Einstein, *Investigations on the theory of the Brownian movement,* ed. with notes by R. Fürth, translated by A. D. Cowper, New York: Dover, (1956).

[2] A. Erdelyi, W. Magnus, F. Oberhettinger, and F.G. Tricomi, *Tables of integral transforms,* McGraw-Hill, (1954).

[3] L. C. Evans, *Partial Differential Equations,* (Second Edition), AMS, (2010).

[4] N. Watson, *A theory of subtemperatures in several variables,* Proc. London Math. Society 26 (1973) 385–417.

[5] B. Fedosov, *Asymptotic formulas for eigenvalues of the Laplacian in a polyhedron.* Doklady Akad. Nauk SSSR, 157 (1964), 536–538.

[6] W. Fulks, *A mean value theorem for the heat equation,* Proc. Amer. Math. Soc. 17(i) (1966) 6–11.

[7] I. S. Gradshteyn and I.M. Ryzhik, *Table of integrals, series, and products,* eighth edition, Elsevier/Academic Press, Amsterdam, (2015).

[8] S. Hawking, *Zeta function regularization of path integrals in curved spacetime.* Comm. Math. Phys. 55 (1977), no. 2, 133–148.

[9] L. Li & A. Strohmaier, *Heat kernel estimates for general boundary problems,* J. Spectral Theory, 6, (2016) 903–919, (arxiv:1604.00784v1).

[10] Z. Lu & J. Rowlett, *The sound of symmetry.* Amer. Math. Monthly 122 (2015), no. 9, 815–835.

[11] Z. Lu & J. Rowlett, *One can hear the corners of a drum.* Bull. Lond. Math. Soc. 48 (2016), no. 1, 85–93.

[12] W. Lück and T. Schick, *L2-torsion of hyperbolic manifolds of finite volume,* Geom. Funct. Anal., 9, no. 3, (1999), 518–567.

[13] R. Mazzeo & J. Rowlett, *A heat trace anomaly on polygons,* Mathematical Proceedings of the Cambridge Philosophical Society, vol. 159 no. 02 (2015) 303–319.

[14] R. Melrose, *The Atiyah-Patodi-Singer Index Theorem,* Research Notes in Mathematics 4. A K Peters, Ltd., (1993).

[15] M. Nursultanov, J. Rowlett, & D. Sher, *How to hear the corners of a drum,* Matrix Annals `https://www.matrix-inst.org.au/2017-matrix-annals/`

[16] B. Osgood, R. Phillips, P. Sarnak, *Extremals of determinants of Laplacians.* J. Funct. Anal. 80 (1988), no. 1, 148–211.

[17] B. Osgood, R. Phillips, P. Sarnak, *Compact isospectral sets of surfaces.* J. Funct. Anal. 80 (1988), no. 1, 212–234.

[18] S. Rosenberg, *The Laplacian on a Riemannian Manifold,* London Math Soc. Student Texts 31, Cambridge University Press, (1997).

[19] M. van den Berg, *Bounds on Green's functions of second-order differential equations,* Journal of Mathematical Physics 22, (1981) 2452–2455.

[20] M. van den Berg & S. Srisatkunarajah, *Heat equation for a region in R2 with polygonal boundary,* J. London Math. Soc. (2) 2 no. 1 (1988) 119–127.

[21] K. Watanabe, *Plane domains which are spectrally determined,* Ann. Global Anal. Geom. 18 , no. 5, (2000), 447–475.

[22] K. Watanabe, *Plane domains which are spectrally determined. II.* J. Inequal. Appl. 7, no. 1, (2002), 25–47.

[23] G. N. Watson, *A treatise on the theory of Bessel functions, Cambridge Mathematical Library,* Cambridge University Press, Cambridge, 1995, Reprint of the second (1944) edition.